

Water Models and Water Politics: Design, Deliberation, and Virtual Accountability

Steven Jackson
School of Information
University of Michigan
301D West Hall, 1085 S University Ave
Ann Arbor, MI 48109-1107
734-764-8058

sjackso@umich.edu

ABSTRACT

Computer simulation models have emerged in recent decades as increasingly prominent technologies within the toolkit of modern democratic governance. Despite and/or because of this centrality, however, formerly ‘technical’ domains of modeling have been opened up to new forms of public debate, scrutiny and critique, with uncertain policy consequences. This paper traces such dynamics through one field of contemporary relevance: the joint evolution of simulation models and water management in California. Rather than decrying the politicization or debasement of expertise, I argue that broadening the deliberative basis of model design and use is likely to improve both the technical and political functioning of models. The paper concludes by sketching a model of ‘virtual accountability’ meant to inform the actions of future model builders, users, and stakeholders in contested realms of public policy.

General Terms

Management, Design, Standardization, Theory, Verification.

Keywords

Simulation, Validation, Models, Water, Environment, Policy, Governance, Infrastructure.

1. INTRODUCTION

Numeric models are as old or older than the practice of modern science and government themselves. They have long played a central role at the three-way interface of theory, data, and action, generating the sorts of hypotheses, predictions, and proofs that have become staple elements within liberal democratic regimes of justification and public action [13, 34]. With the emergence of digital computing, the role of models in science and administrative decision-making has taken a qualitative step forward. From physics and the bomb, computer simulation techniques have spread rapidly through science, industry, and public policy in the post-war period, most notably in well-funded

domains where theoretical, observational, and experimental roads to knowledge have proven impractical, whether for reasons of cost, accessibility, or ethical sensitivity. These have included the big (e.g. climate science), the small (e.g. genetic research), the remote (e.g. astrophysics, mineral exploration), the vastly distributed (e.g. epidemiology, economics), and the humanly, ecologically, or politically fragile (e.g. medical research, post-proliferation nuclear weapons testing) [12, 16, 17, 23, 31, 32].

Despite this remarkable growth, there has been as yet little work addressing simulation models as *policy technologies*, i.e. artifacts whose technical shape, deliberative character, and public effectiveness are built and adapted through mutually iterative processes of technical and political refinement. To understand this dynamic, digital government scholarship must find answers to a number of prior questions. What role have models come to play in mediating the deep political and epistemic tensions characterizing such fields as environmental management and policy? What barriers and limitations have models encountered in their movement from academic and agency science into the world of public deliberation and decision-making? How have the distinctive characteristics of that world acted back upon the technical work of simulation, influencing the design histories and affordances of the models themselves? Finally, what prescriptive lessons can be drawn from these experiences to guide the future work of model builders, users, and stakeholders in contested realms of public policy?

Drawing on scholarship in Science and Technology Studies (STS) and the methods of the qualitative social sciences, this paper addresses such questions through the medium of a single case study: the technical and political evolution of water simulation models in California and the American Southwest, and their shifting place within the region’s high-stakes and often fractious water management regime.¹

¹ Principal fieldwork for this paper was carried out between January 2003 and July 2005, and included ethnographic observation and more than 60 interviews with modelers, policymakers, and environmental, urban, and agricultural representatives. For a more complete description of study methodology and findings, see [18].

2. MODELS IN ECOSYSTEM SCIENCE AND POLICY

Among the policy fields in which computer modeling has assumed a prominent post-war role, environmental science and management have represented something of a perfect storm: they have tended to draw on vast amounts of data sourced from highly distributed collection regimes; they deal regularly in large, disparate, and complexly interconnected data sets; they occupy research fields in which the possibilities of experiment or direct observation are prescribed by constraints of time, scale, safety, or physical access; and they have grown in partial response to social movements and agencies with a strong pragmatic interest in prediction [12, 30]. Under such circumstances, models have come to play significant and indeed multiple roles: from the shaping of observation and experiment, to the synthesis of large and disparate data sets, to the generation of numeric predictions and forecasts [6]. In each case, model builders and users have faced significant challenges. Modeling work has frequently evolved in relative isolation from other analytic and fieldwork traditions, leaving key model parameters, data sets, and functional relationships incomplete, ill-documented, unevenly maintained, or otherwise underspecified [2, 10, 24]. Modelers in many fields have struggled to define appropriate relationships between data sets of radically varying types and quality, especially where real-world functional dynamics – often the question of central analytic concern – remain poorly understood. Additional theoretical and practical challenges greet efforts to define the appropriate scope and granularity of simulation: how much of the (impossibly vast) world ought to be accounted for within the confines of the model, and what can be safely left outside of it? At what point does the radical and fine-grained inclusion of variables obscure the analytic clarity, general comprehensibility, and pragmatic workability of models? Here, modelers have debated the relative merits of ‘fast-and-frugal’ models (marked by speed, agility, access, and rapid evolution) versus their larger, slower, and more computationally intensive relatives [7].

Modelers have adopted a variety of strategies for dealing with such uncertainties. A first and obvious response has been to measure model results against real world experience, adducing direct or indirect evidence for and against the reliability of model results. But this check against external record, simple in concept, may be surprisingly difficult to accomplish in practice – not least because models are frequently deployed precisely where possibilities of observational or experimental knowledge are at their weakest. Under such circumstances, ‘crucial tests’ able to conclusively confirm or disconfirm model results are in remarkably short supply.

In response, model builders and users have turned to a variety of second-order strategies to support the validity of their claims. In sensitivity analyses, modelers tweak variables in a sequential manner, measuring the degree of responsiveness or “sensitivity” registered across key parameters of the system. Sensitivities deemed to be inordinate when measured in this way may indicate the presence of artifactual properties likely to skew model results vis-à-vis the real-world systems under study. A second response to uncertainty in simulation has been the historical calibration or validation exercise, where the model is evaluated according to its ability to reproduce known results captured from historical observation. Under the simplest description of this, the model is

loaded with the initial conditions and known input data for a given period, run, and the results compared to the historical record. Alternatively, the period of calibration may be split into two separate phases: an initial ‘tuning’ period, in which model parameters are adjusted until a good fit with historical data is achieved; and a second testing phase, in which the corrected model is run and compared against the remainder of the record. Models capable of reproducing known history with reasonable accuracy are inferred to reliably mimic and project the performance of the system into the future. Models tested in this way are frequently said to be “verified” or “validated.”

Such second-order responses to uncertainty face problems of their own, however. As Oreskes et. al. point out, in contrast to the possibility of closure existing in formal logic and mathematical systems, the ‘systems’ of ecosystem science and management remain radically and inevitably open: input parameters are incompletely known, model elements remain subject to uncertain scale effects, available data is of uneven quality and coverage, and the fit of model to world depends on a prior set of ultimately untested and unmodeled assumptions [1, 27]. In practice, efforts at historical validation as described above are often in a strict sense inconclusive, since model tuning constitutes an ordinary and ongoing part of model life, limiting the degree of autonomous verification that may be said to follow from successful replication of the historical record. Even where the historical match is good, the frequent non-linear character of earth systems, when combined with the generally short periods and places for which good historical data is available (as compared to the scales of time and space at which models are often called upon to predict) raises the problem that currently negligible mismatches between model and world may lead to significant divergences over time: models only marginally wrong over the recent past may prove to be significantly, even catastrophically, wrong when projected far enough into the future. Moreover, successful calibration does not necessarily imply a ‘true’ or ‘realistic’ understanding of underlying causal mechanisms and dynamic; models can produce the ‘right’ numbers for the wrong reasons, and the significance of fundamental conceptual error may increase as the system moves downstream in time. Beyond all this, the fundamental openness of ecological systems means there is no guarantee against future changes in the system that might render models reasonably accurate in replicating past and present observational results wildly inaccurate when projected into the future. Under such circumstances, as Oreskes and Belitz observe, efforts to train simulations to historical data may exacerbate the conservative bias of models by extending, sometimes without justification, current trends into the future. Similar dynamics govern the general exclusion of low probability events, whose collective impact may further skew model results [28].

Such narrowly ‘technical’ challenges take on new life and complexity as they travel beyond the immediate boundaries of professional communities of practice. An important impetus for the development of specific models and the growth of simulation technique more generally has come from agencies and policy-makers – themselves driven by changes in the political and regulatory field wrought in large part by various environmental movements – who have sought in models new and authoritative bases on which to ground regulatory and policy actions. In this regard, modelers work in a field simultaneously constrained and constituted through the presence of what historian of science Ted

Porter has described as ‘powerful outsiders’: figures from beyond the immediately technical realm who nevertheless exert a significant influence over the shape and form of its internal deliberations and practices [29]. One effect of this positioning vis-à-vis the decision requirements of the policy realm may be to push models in the direction of a steadily harder predictive stance, as sharply bounded simulation results pass into the hands of actors for whom technical caveats, uncertainties, and other limitations will remain at least opaque and perhaps an impediment to efficient decision-making. Under such circumstances, predictions may take on an aura of finality that makes even their producers uneasy. The fallibility or fragility of model knowledge may disappear in translation, peeled away on the margins of the science-policy interface. By virtue of such dynamics, model predictions, like other scientific findings, will frequently appear to ‘harden’ as they travel outwards; in Harry Collins memorable phrase: “distance lends enchantment” [9]. Where such dynamics are most fully developed (e.g. the climate change debates), this has tended to produce an all-or-nothing public response to model credibility: models are too frequently regarded as all right or all wrong, with little room for a nuanced and ultimately more helpful engagement with model strengths and weaknesses.

The upshot of all this is that the degree of certainty both professionally and popularly assigned to model predictions may be systematically distorted. As Oreskes et. al. note, routine terms of art such as ‘verification’ and ‘validation’ may mislead, particularly where restricted professional usages meet the more expansive and unqualified meanings generally assigned these terms in common discourse [27]. Under such conditions, assessments of fit and adequacy remain deeply situated exercises. Competent and accredited professionals may well (and frequently do) arrive at different conclusions as to the validity and appropriateness of a given model application, even where using the same data. External stakeholders (e.g. funders, government agencies, scientific review panels, industry officials, public interest groups, etc.) will bring criteria of fit all their own, which may contradict, but also significantly shape, the internal deliberations of model builders and users. To the extent that models (and model results) function as ‘boundary objects’ shared between multiple social and institutional worlds, assessing their adequacy is an inescapably practical as well as narrowly technical affair [31, 33]. To the extent that models are purposeful, i.e. embedding particular predilections or orientations to action, questions of design are inextricably bound with questions of application. In this context, the central question of model evaluation is not ‘is it good?’ but rather “is it good enough for the purpose?” [15].

Under such circumstances, the contributions of modeling to wider arenas of public debate should be offered and received in a spirit of responsible and critical humility [19, 20]. At the end of the day, as Oreskes and Belitz put it, “the most we can do is to say that a model is close to the state of the art (if it is), that it has been grounded in our best understanding of known natural processes (if it has), and that we built it on the basis of abundant, well-constrained empirical input (if we did)” [28]. The modes of a mature public encounter following from such an admission – what I address under the language of ‘virtual accountability’ – remain very much to be worked out.

3. MODELS AND WATER MANAGEMENT: A BRIEF HISTORY

Computer simulation models made their first direct appearance on the California waterscape beginning in the 1960s, appearing more or less simultaneously within the planning and operations divisions of the California Department of Water Resources (DWR). On the operations side, the build-out of the managed water network, with the addition of the State Water Project to the now nearly complete Central Valley Project, led to new technical and organizational challenges around the coordination of an increasingly complex and inter-tied network. By wiring the reservoirs in real time and developing computational models and monitoring protocols around the resulting data, project operators could more safely manage the growing complexities of a multiply inter-tied system. At around the same time, DWR and Bureau of Reclamation planners began developing a series of depletion and accretion studies simulating the effects of varying hydrological conditions on streamflow and groundwater patterns in the all-important Central Valley.

By the early to mid 1980s, these fledgling efforts had grown and achieved a measure of standardization. Efforts to link facilities connecting through the State Water and Central Valley Projects were largely complete, with real-time management information now converging on the Operations Control Center in Sacramento. Initial planning studies commissioned on a one-off basis had morphed into two more general modeling frameworks: DWRSIM, used by the California Department of Water Resources to manage the State Water Project; and PROSIM, used by the Bureau of Reclamation in its Central Valley operations. Beyond the state and federal models, many water agencies, irrigation districts, municipal suppliers and academic modelers developed simulations of their own. Some of these operated in loose coordination with the state and/or federal modeling efforts; many others were designed and run in isolation.

By the early 1990s, this heteronomy of models and modelers had become a source of both technical and political instability. The resolution of discrepancies between the state and federal models absorbed increasing amounts of time and institutional resources. A series of disputes between the DWR and local water agencies resulted in acrimonious hearings before the state water board in which modelers on either side lined up to challenge the integrity and credibility of the opposing model. Such publicly-aired rivalries within the still loosely defined professional circles of modeling gave an increasingly restive set of external actors new purchase on the previously opaque details of water management and engineering; environmentalists, agriculturalists and urban users opposed to a particular policy stance adopted by the DWR or Bureau of Reclamation were in some cases able to point to other numbers and alternative conclusions reached by different sets of formally accredited and apparently equally legitimate analysts.

At the same time, under the weight of the new technical, political and institutional demands placed upon them, the now legacy models of DWRSIM and PROSIM were beginning to break down. At the level of design, the ‘spaghetti-coded’ and idiosyncratically-produced nature of this generation of models meant that few actors other than their original creators could be said to understand their operation in enough detail to recognize and correct the frequent errors and inconsistencies that emerged in

operation. This posed, among other things, a unique personnel problem: as original modelers left the DWR or Bureau for the lucrative engineering consulting industry, government agencies found themselves in the embarrassing position of being unable to understand and run their own models (short of hiring back their own former employees at private consulting rates). At the same time, the ‘hard-wired’ nature of the models (i.e. the need to specify elements and perform changes at the level of opaque, usually Fortran, code) made comparative analyses undertaken on the basis of models both awkward and time-intensive.

Significantly, even such apparently ‘technical’ problems were embedded within and substantially owed to a wider set of institutional, political, and ecological transformations. For instance, the comparison problems posed by the hard-wired nature of DWRSIM and PROSIM could be overcome, or at least accommodated, within the relatively stable management and policy regimes inherited from the 1950s and 60s. Under such circumstances, the demands placed on the model were relatively simple: maximize deliveries and give some consideration to power generation, subject only to the constraints of flood control. This situation changed drastically as ‘environmental’ claims on the system mounted, and in particular following the passage and enforcement of the National Environmental Protection Act (NEPA) and California Environmental Quality Act (CEQA), which made Environmental Impact Reviews part of the language and responsibility of water managers throughout the state. Now, suddenly, comparative analysis was mandated by law (backed by activist pressure), and the previously only dimly experienced ‘technical’ inadequacies of the models were made glaringly apparent. Relatedly, the problem of spaghetti-coding, noted above as a personnel issue, reemerged under the pressures of the environmental movement as a fundamental problem of legitimacy: what faith should public actors place in a model whose inner workings remained opaque to all but the smallest handful of the initiated, especially in cases where such models provided the primary, even the sole, evidentiary basis for public decision-making? [18]

3.1 CalSim: consensus and controversy

Faced with such pressures, the design, practice, and politics of water models in California underwent a series of important changes in the early to mid- 1990s. Most importantly, traditional rivals DWR and BR agreed to cooperate in the joint development of CalSim, a new and widely-touted “consensus model” that would replace each of their aging proprietary models. At the level of technical design, CalSim would correct many of the shortcomings noted in its predecessors: it would be soft-wired (or “data driven”) as opposed to hard-wired, lending itself more readily to the sorts of comparative and speculative studies its predecessors were ill-equipped to handle; it would follow the now-common coding principles of structured and object-oriented programming, allowing improvements in general readability and new possibilities for modular development; it would incorporate new and standardized user interface and file management procedures that would ease the flow of data in and out of the model; it would be open source and, in principle, freely downloadable from the DWR website; and it would pull all representations of data, including the model’s crucial operating rules and assumptions, from FORTRAN code to a more flexible and accessible natural language interface, keyed specifically to the conditions and practices of western water management.

Collectively, it was argued, such changes would go a long way towards repairing the technical consensus around water management fractured by the disputes and discrepancies of the eighties and early nineties. Converging on a common model, it was also suggested, would realize new efficiencies of scale by coordinating the development efforts of DWR, BR and other modelers around a single common object. As an object shared between agencies, CalSim would grow faster, more efficiently, and more reliably [8, 11, 14].

Standing next to and supporting such technical claims were a series of explicitly political appeals. The newly open architecture of CalSim, it was suggested, could contribute importantly to public confidence in the tool, establishing a form of political legitimacy that its predecessors had at first not needed, and then distinctly lacked. With this common understanding in place, it was hoped, a substantial portion of the legal and political controversies that had embroiled the system since the rise of the environmental movement could be done away with, and the various parties (but most especially the managers in the DWR and Bureau) could get back to the work of the rational management and distribution of resources. In this regard, CalSim was touted as the putative lynch-pin in a peace-through-science settlement promising to restore both order and a measure of civility on the California water system.

Underlying all of these hopes lay the common techno-political dream of *transparency*. Through such innovations as structured programming, natural language interfaces, standardized file management procedures, and modular (object-oriented) development strategies, CalSim aspired to a level of architectural transparency far surpassing the opaque code of its predecessors. Such innovations, it was argued, would improve model reliability, supportability, and cross-agency technical collaboration. But if transparency was an architectural virtue, it was also a democratic one: in the fractious climate of California water politics in the early 1990s, the very ‘openness’ of the model – open source, open code, open documentation – was heralded as an important and necessary *political* accomplishment. On both the architectural and democratic fronts, however, transparency was a partial and tenuous achievement at best, consistently undermined by the challenges of data and organizational alignment and the in some ways irreducible complexity of the system under representation. More subtly, the principle of transparency, while commonly presented as a solution to the absence of trusted relations, turned out to depend on them. From this perspective, the ‘external’ transparency of models rested substantially on the ‘internal’ stabilization of its constituent parts through the principled agreement to leave certain assumptions and assertions (including the professional competence and good faith of its practitioners) unquestioned. Absent this level of stabilization, as subsequent events would show, the promise of transparency failed and models could be rendered once again vulnerable to critique, dispute, and the real-world machinations of western water politics [18].

4.0 THE TROUBLE WITH NUMBERS

4.1 The State Water Project Reliability Report

In August of 2002, the California Department of Water Resources released a draft document with the apparently innocuous title of “The State Water Project Delivery Reliability Report.” The task of the report was seemingly straightforward: to “provide current information on the ability of the SWP to deliver water under

existing and future levels of development, assuming historical patterns of precipitation” [3]. This exercise in predictive modeling, undertaken using CalSim and the 73 years of annualized data contained within the acknowledged period of record, was complicated by two contextual factors. First, the report itself grew out of (and was indeed mandated by) ongoing legal controversies surrounding the 1995 Monterey Amendments to the State Water Project contracts, which environmental groups charged with fundamentally over-stating the delivery capacities of the state water system – the aptly-named problem of ‘paper water’ – thereby throwing open the door to unrestricted and unsustainable growth in the state. Second, recently passed bills in the California Assembly requiring private land developers and local planners to demonstrate water supply reliability twenty years into the future had essentially granted CalSim – the only tool deemed capable of making this sort of prediction – the weight of law. Within this politically-charged climate, reactions to the DWR’s projections were swift. In written and verbal testimony submitted as part of the Report’s public review and comment period, the modeling on which the analysis was based was attacked as deficient on a number of grounds: for failing to account for the potentially serious effects of climate change on regional water supplies; for its inadequate attention to water rights senior to the State Water Project (including native, municipal, count-of-origin, and public trust claims) which could limit future deliveries through the state system; its insufficient representation of dynamics within both the federal portion of the system and the conjoined groundwater system; and for its decision to hold regulatory constraints on the system constant, thus failing to account for the likelihood of either future infrastructural development that would increase project supply capacity, or future endangered species claims that would effectively reduce it.

Arguably, none of these detailed and somewhat arcane technical debates would have entered the public sphere at all, were it not for two of the report’s central, and deeply counter-intuitive, findings: first, that delivery reliability would actually *improve* over the course of the 2001-2021 period, in spite of the increasing upstream demands placed on the system; and second, that the SWP could be relied upon (at both 2001 and 2021 levels of development) to deliver water at levels that were, on average, nearly *fifty percent higher* than historic deliveries. In contrast to real-world SWP deliveries hovering in the neighborhood of slightly more than 2.0 million acre-feet (maf) per year, the report announced simulated deliveries ranging, on average, from 2.96 to 3.13 maf (at 2001 and 2021 levels of development, respectively). By what logic, asked the report’s many critics, might the constraints on an already overtaxed and still tightening system be expected to *ease* over the next twenty years, at a moment when virtually everyone in the California water community was predicting and preparing for a much darker scenario of growth, shortage, and conflict? As one critic noted, somewhat incredulously, “We are asked to believe that the SWP will reliably, on average, provide an additional million acre feet of water (50% greater than past performance). The finding defies logic and is inconsistent with the system’s actual performance.” [35].

These questions took on added political weight when Senator Michael Machado, member of the powerful Senate Committee on Agriculture and Water Resources, wrote to express his concerns. Noting the widespread public backlash and tension surrounding

the draft document and the modeling underlying it, Machado argued that the report was ‘premature’ and urged DWR to take active steps to address the concerns and criticisms leveled against it. As noted by Machado, the stakes went well beyond the report itself:

- Local development could be hampered if, when complying with SB 221 (Kuehl) and SB 610 (Costa) of California’s Environmental Quality Act (CEQA), there are significant disputes over current and future water supplies.
- Conclusions of CalFed’s Integrated Storage Investigations (ISI) will be suspect given that the same model is used in both the ISI and the SWP Reliability reports.
- Future statewide bonds for increasing water supply will be in jeopardy, if opponents can credibly challenge the underlying analysis.

As a measure of his concern, Machado took the unusual step of asking the California Research Bureau, the research arm of the state library system, to produce a formal analysis and comment on the report. The CRB commentary, drafted by Assistant Director Dennis O’Connor, constituted the longest and most detailed intervention over the course of the SWP reliability report controversy, weighing – and in part endorsing – the claims of the report’s environmental critics² [26].

Dominating the public hearings and comment period, as feared by Machado and detailed in the CRB report, were public concerns and criticisms around CalSim itself. These were organized around two interlinked questions: first, the manifest (but underacknowledged) limitations of the model in use, as shown up in the specific context of the report’s production; and second, the credibility of models as policy tools more generally, particularly where supplying the primary evidentiary input to public decision-making. On the first point, critics were quick to point out the remarkable poverty of the model when it came to representing important facets of the California water system – most notably, surface-groundwater interactions and delivery curtailments following unacceptable takes of endangered species at project pumping facilities – that must play a significant role in all future projections of real-world supply availabilities. Similarly, critics noted, CalSim assumed a level of order and predictability in the *human* operation of the system that was not always in evidence; complex operational rules and legal infrastructure aside, the various actors in the system, from DWR operators to project contractors, did not always follow the logics and rationalities that the model (though in principle, also their operating and contractual obligations) imposed upon them. This was particularly

² Specifically, the CRB report questions the Department’s suggestion that the reason historical deliveries fall well below modeled results is that the contractors haven’t requested their full allotments in past; the apparent failure to account for the effects of upstream development (and thus consumptive use) under the 2021 scenarios; the weak representation of key variables such as groundwater interactions; the discrepancy between CalSim monthly results and the more modest projections associated with finer-grained daily models; and the use of the model for predictive rather than comparative purposes.

true when the stakes were highest (most notably, under drought conditions) when operators and other actors were required to cobble together local responses to crisis that strayed from and sometimes violated the formal operating procedures laid down in the model. For all these reasons, CalSim was too simple, too narrowly framed, and entirely too thin to incorporate the degree of nuance and complexity needed to reliably project the real-world futures of the system.

Beyond this, critics felt that the DWR was essentially asking them to accept *on faith* the efficacy of a rather opaque technological artifact to which they had been granted little effective access, let alone participation. Critique returned time and again to the fact that at the time of the report, CalSim had yet to undergo anything like the sort of rigorous external assessment that might establish some grounds for its legitimacy, despite the fact that it had then been in use for several years by analysts within the department, the Bureau of Reclamation, and a number of informally affiliated private consulting firms. Critic after critic noted that CalSim had yet to be tested against the historical record in anything like the sort of validation or calibration exercise typically expected of models in other domains of science and public life. Without knowing that the model could effectively reproduce the intensely-observed history of water in California, why should those suspicious of its claims (and the motivation of those making them) grant their assent? Similarly, critics noted, CalSim had never been reviewed in anything like a systematic way by anyone other than its creators and principle users within the department, bureau and a few hand-picked consultants. Given its centrality to water planning and management in the state, why hadn't a formal peer review been conducted? In the absence of these initiatives and/or the presumption of good faith on the part of the department, there appeared to be little left to compel general assent to the model and its claims. In this context, as one critic noted sardonically, "Our model says so' is not enough to base policy on" [35].

By the end of the SWP reliability report public comment period, the concern cited in Senator Machado's letter (and privately conceded by modelers and planners in the Department of Water Resources) had been borne out: passing more or less quickly over the details of the report, public controversy had come to rest squarely on the credibility of CalSim itself. In the process, previously arcane details of model design and operation were discussed, debated, and sometimes challenged by actors well beyond the usual core of modelers, engineers and departmental managers. Such events contributed to a larger mood of public skepticism that disrupted and arguably put to rest the hopes associated with the peace-through-science settlement of the 1990s.

4.2 *The 2005 Water Plan Update*

While concerns around CalSim and the credibility of models more generally were surfacing in the context of the SWP Reliability Report, models were being opened up from a different direction in the context of statewide water planning. The immediate venue for this debate was the California Water Plan, a once every five year attempt to square the circle of California water policy, synthesizing the starkly different interests of agricultural, urban, environmental, and other policy claimants into a credible and workable statewide management framework.

The most recent water plan update process, which issued its draft report for public comment in summer 2005 (a full two years

behind its legally mandated schedule) grew out of a particularly acrimonious set of debates culminating in the bitterly divisive water plans of 1993 and 1998. As the record of public comment reveals, commentators on the earlier plans differed sharply as to the nature of the crisis unfolding in California. For a range of environmental critics, planning was urgently required to redress the great ecological disaster long unfolding in the Bay-Delta and other areas of the state. For agricultural supporters and some urban water agencies, the crisis was precisely reversed, namely, that the accumulated weight of population growth and environmental demand had rendered the water supply system fundamentally unreliable and in particular vulnerable to future fluctuations of the hydrological cycle. Not surprisingly, respondents also varied in their estimation as to where the elusive 'new water' needed to fix the California system would come from: urban and agricultural contractors in the state looked for the most part to new surface storage and conveyance facilities, while environmental advocates argued in favor of so-called 'soft path' strategies, in the form of efficiency gains, strict conservation measures, and other sorts of demand-side management strategies [36]. Complicating all these debates were the looming effects of regional climate change, feared by many to reduce the natural system's capacity to carry over winter precipitation into the peak summer months of agricultural, urban, and power consumption. [21].

These varied policy positionings, which cut to the heart of the future growth and development strategies of the state, returned time and again to apparently technical disputes around the description, quantification, and prediction of water. Scientific understanding of groundwater depletion and recharge processes was appallingly bad, charged some commentators, leading to overdraft figures, projections, and overall water balances that were little better than guesses. The plans substantially overestimated future agricultural demands, charged others, underplaying the effects of future efficiency gains and the water-saving potential of continuing market-driven 'ag-to-urban' transfers. The plans similarly exaggerated future urban uses, others argued, which were based on demographic projections that failed to consider the potential dampening variables of economic recession, land price inflation, etc. In both cases, the failure to assign real-cost pricing – i.e. modulating demand projections according to the rising prices that would (or should, in the absence of ongoing subsidies to agriculture) accompany future water scarcities – significantly skewed both urban and especially agricultural demand in an upward direction. Estimates of present and future 'environmental water' needs – counted as a line item for the first time in the 1993 update – were argued to be either too high or too low, and in any case inadequately specified and/or based on a level of scientific understanding insufficient to justify the large-scale restructuring of project deliveries [18].

It was against this acrimonious backdrop that work on the current water plan began. Vowing to redress the participatory failings of 1998, DWR retained professional mediation services and expanded the plan's official Advisory Committee to 65 people, including for the first time district level, tribal, and environmental justice representatives. At the level of content, three consequential decisions were taken. First, in an effort to acknowledge the deep uncertainties facing water prediction and management in the state, the controversial single figure "gap analysis" of past plans (subtracting current supplies from projected needs, and proposing

facility or management changes to redress the balance) was dropped in favor of a 'scenarios' approach, in which the performance of the system under multiple future supply and demand conditions would be contemplated. Second, in the face of widespread skepticism surrounding the Department's procedures for normalizing data into water year 'types', the advisory committee elected to work from real data sourced from three recent water years: 1998 (classed as a wet year), 2000 ("average"), and 2001 (the driest on record since the 1987-1992 drought). Arguably the most significant development, however, came with the decision to abandon numbers altogether – or rather, to put off the thorough processing of them until later stages of the plan. In a sharp departure from prior plans, Bulletin 160-03 would be issued sequentially, with a policy-focused first phase describing the current system state and describing general priorities and potential policy stances, a second tool-building phase establishing in more detail the precise approaches to be taken in quantifying the system, and a third phase in which the qualitative scenarios outlined in phase one would be populated with data and at last calculated out. There were some immediately pragmatic reasons for adopting this strategy. By early 2003, the plan was far off its timeline and showed little or no hope of hitting its scheduled release date at the end of the year; the turnover of senior personnel within the Department following the gubernatorial recall election of 2003 had recently introduced additional delays and uncertainties. There was also some sense, shared among DWR officials and advisory committee members involved in the planning process, that putting off the zero-sum game of calculation, like the scenario decision before it, may have softened the sharper edges of interest group conflict and therefore played a role in keeping stakeholders committed and engaged in the planning process.

At the most basic level, however, the decision to prepare and release the plan in stages was owed to widespread reservations around the quality and trustworthiness of numbers. Through the early stages of planning, the broad lines of quantitative disagreement, like the political split more generally, followed those laid down in the aftermath of the 1998 plan. Environmental groups contested DWR procedures for calculating urban and agricultural demand (in particular, its utter neglect of price signals), pushed for new ways of calculating the savings to be achieved through urban and agricultural efficiency, and urged the state to adopt beneficiary-pays and true-cost pricing principles. Agricultural groups and urban suppliers argued that the numbers on "new water" produced through urban and agricultural efficiency improvements were wildly optimistic, and suggested that many of the 'soft' gains to be had by such measures had already been achieved during the 1987-1992 drought (the so-called 'demand hardening' argument). Agricultural representatives argued further that such projections implicitly endorsed an expanded and ultimately short-sighted program of ag-to-urban water transfers that would leave the state unable to meet its own 'food and fiber' needs within the foreseeable future. By 2003, as the plan's official delivery date neared, numeric and political tensions on the advisory committee deepened. Following early efforts to avoid the traditional sectional splits into farm, city, and environment – as one respondent later noted, "we were trying at the start to not go positional" – around the middle of 2003 the committee regrouped itself into caucuses, with representatives now speaking on behalf of the traditionally-identified groups.

Through this process of asserting and disputing the adequacy or otherwise of specific numbers, the advisory committee gradually came to a more general awareness of the limits and problems of data and models in general. In 2002, following widespread expressions of concern within the advisory committee over the credibility of the models on which the 2030 projections were to be based, several members of the advisory committee formed a Modeling Work Group, dedicated to the task of exploring and reporting back to the group on the strengths and limitations of available data and modeling frameworks. In September 2002, the group prepared a formal modeling proposal which was subsequently adopted by the advisory committee. In contrast to the technocratic certainty characterizing the language of previous water plans, the advisory committee statement struck a pointedly skeptical note. While acknowledging that the "proposed models have some constructive role to play in Update 2003," the work group cautioned that "the potential exists for policy makers and the legislature to misuse modeling data, which necessitates judgment in releasing select results and identifying model limitations" [5]. Beyond this,

Models are inherently uncertain. Any decisions based on models should include this caveat. All models in Update 2003 have limits: DWR staff and the Advisory Committee will identify those limits in the main plan's text and provide details in the appendix. The Advisory Committee will bear such limits in mind and reflect on improvements when interpreting the results of model runs. [5]

Far from a ringing endorsement or a blanket condemnation, the response of the advisory committee to the presentations of DWR modelers was both critical and pragmatic. The members of the committee (and in particular its modeling work group) were willing to acknowledge the usefulness of models as a potential input to policy-making, but were not willing to grant their assent on faith, or to cede to model results the preponderant weight in future water decisions. In the face of such ongoing uncertainty, advisory committee members were urged to rely on their "collective wisdom," and treat the predictive claims of the models with a degree of informed skepticism.

Such skepticism also became the occasion for a fundamental rethinking and the beginnings of a redesign of the state's modeling infrastructure. As members of the workgroup noted, the existing suite of models and numeric analysis tools was significantly, perhaps even dangerously, misaligned with the sorts of questions water managers and public decision-makers in the state were increasingly being called upon to address. As one advisory committee member noted in a letter to the committee,

Most planning analysis and data collection for California's statewide water resources were developed for an era of large-scale water facility development. Our analysis capability continues to specialize in the operation and planning of the large State and Federal water projects. Most analysis largely neglects the local and regional activities which are the hallmark of current water management, such as water conservation, conjunctive use of

ground and surface waters, water transfers, and wastewater reuse. [25]

On this basis, “DWR’s data collection and analysis capabilities must be substantially re-directed and re-engineered to re-orient DWR planning to aid, support, and integrate local and regional efforts.” [25]

In the end, the perceived weakness of the available numbers and models led the advisory committee and DWR planners to the three-phase approach noted above: they would produce a plan, but it would, at least in the interim, contain very few numbers, and certainly none of the summative sorts of numbers associated with things like the reviled ‘gap analysis’ of past plans. At the same time, work would begin on a second phase, in which current data and model deficiencies would be identified, and long-range approaches to correcting these undertaken. Armed with the new numbers and tools, the plan’s third phase would at last ‘cost out’ numerically the scenarios generated qualitatively in phase one. The draft of the long-awaited plan’s first phase was released for public comment during the summer of 2005 (where, predictably enough, its lack of numbers came up for regular criticism). The third and final phase is now projected to arrive in 2008 – five years late, and in precisely the year the *next* water plan was to have been delivered.

5.0 VIRTUAL ACCOUNTABILITY

The efforts of the water plan advisory committee and its modeling workgroup represent only one part of a larger effort to address and restore the credibility of models and modeling as input to public policy that has been arguably damaged by recent controversies in California water management. Stung by criticisms received during the Monterey Amendment / SWP Reliability Report controversies, the DWR, acting in conjunction with the Bureau of Reclamation, has taken several steps to address at least the most immediate concerns of its critics. In 2003, the DWR issued the findings of its Historical Operations Study, the most serious attempt to validate CalSim vis-à-vis the historical record to date. Running the model against operating logics, streamflow data, and water delivery records from 1975-1988, the study authors argued that CalSim in fact performed remarkably well, returning Delta outflow figures that differed on average by only 7% from historical values, and hitting within 5% during the crucial drought period of 1987-1992. On the question of groundwater, where hard data was (and remains) notably lacking, CalSim was tested against the more detailed representation contained within the Central Valley Groundwater Surface Water Model and found to be broadly, though not perfectly, compatible [4]. In March 2005, DWR officials presented preliminary results from the first large-scale CalSim sensitivity analysis, in which the model again performed reasonably, though not perfectly, well.

In arguably the most significant development to date, in 2003, the CALFED Science Program (responding to DWR requests) undertook the first large-scale (if still limited) peer review of CalSim, conducted by seven ‘external’ experts chosen for their long experience in water operations and simulation modeling. The results of the review were mixed: while endorsing the overall technical soundness and general approach of the model (praising in particular CalSim’s open source and consensus-building ambitions), a wide range of reservations were expressed: geographic coverage in the model was weak and notably incomplete, in particular with regard to Southern California and

Colorado River transfers; attention to questions beyond traditional supply concerns was weak or non-existent; the distributed character of model development within and beyond the agencies raised important questions of versioning, consistency, and quality control; the model’s understanding of real-world operational dynamics and decision-making was grossly and unacceptably simplified; and despite apparently real and laudable aspirations towards openness, insufficient effort (in the form of documentation, user-friendly interfaces, user support, public workshops, etc.) had been devoted to making the model usable or even comprehensible beyond the confines of a fairly narrow circle of experts [8, 14].

As this list of design responses to controversy (and their shortcomings) may begin to suggest, the real-world challenges of ‘modeling democratically’ within realms of complex and bitterly contested public policy are immense. Moreover, they spill regularly and confusingly beyond the confines of ‘straight’ technical practice into broadly sociological registers of trust, confidence, and credibility which modelers and water managers are ill-equipped by training to deal with (though many have gained considerable practical skill in this regard). Under such circumstances, technical ‘fixes’ to political problems are likely to fall short of their goals (as indeed are ‘sociological’ responses to hard technical concerns). The challenge, as always, is to work across the two sides of this divide simultaneously.

What might an appropriately deliberative solution under such circumstances – what I’m describing under the language of ‘virtual accountability’ – entail? First, it should be noted that other modeling frameworks, some more supportive of stakeholder deliberation, may be identified. ‘Gaming’ models have been developed and successfully deployed in several instances as an aid and heuristic to contentious group decision-making processes. ‘Screening’ models (including periodic calls for a ‘CalSim-lite’) may be developed and deployed in forms that sacrifice a degree of analytic precision and granularity, but may gain in broader stakeholder accessibility and general analytic wieldiness. Neither of these approaches could entirely supplant the multiple functions CalSim is currently called upon to perform; but they could perform at least some of those functions in a more deliberatively-supportive and ultimately effective manner.

Second, as noted by respondents to the CalSim peer review, modelers could also do more to build effective public access to their tools and findings, through better documentation, interface development, public training and information sessions, and other sorts of mechanisms. Some of these are already underway, e.g. regional review workshops fielding public comment around the representation of particular system components within the larger CalSim architecture. This sort of public investment, in what we might think of as ‘translation goods,’ ought to be undertaken as democratic as well as purely analytic investments – a point that funding realities and structures within the DWR are currently ill-equipped to accommodate.³

³ Notably, funding for CalSim is organized under the ‘project’ side of the DWR organizational hierarchy, concerned primarily with the planning and operation of the State Water Project and only secondarily with the wider concerns around water policy and participation traditionally housed in the Department’s planning division.

Third, as painful and inefficient as it may sometimes be to move beyond the measured world of technical decision-making, serious and sustained efforts at broad public engagement seem the most promising road to the longer-term goals of widespread model literacy and trust that in the end will be needed to sustain and extend the viability of modeling as a policy technology. Important early movements in this direction may be identified in the work of the water plan advisory committee. Until California's water problems go away and/or simulation modeling achieves an unquestioned sophistication and place within the pantheon of credible policy knowledges – neither of which seems likely to happen anytime soon – ongoing efforts to engage across the technical-public divide remain the most likely long-term strategy for building trust, confidence, and legitimacy – and ultimately an effective and democratically sensible water policy.

Beyond their immediate implications for simulation modeling in the water policy context, the story sketched above speaks to issues of wide and growing digital government concern. While the entrenched conflicts, high stakes, and deep uncertainties dominating California water policy may make this an unusually intense laboratory for the observation of simulation in action, it is by other measures an old and entirely unremarkable story: actors on all sides of the California water debate are by now well-versed in the challenges of performing technical work in a politically fraught field – the ever present challenge of what I've described elsewhere as 'doing hard politics with soft numbers' [18]. This common tension and dynamic has yet to receive the theoretical elaboration it deserves, in either the digital government or more general public policy literatures.

There are also evolutionary dynamics worthy of digital government and broader public policy note. In each of the controversies sketched above, the apparently technical art of modeling was opened, however awkwardly and painfully, to review and criticism beyond its immediate circle of expertise. In the process, public light was cast into areas of practice formerly ceded almost entirely on trust to a domain of professional expertise. This occurred not through any abstract notion of participation or transparency (though it has clearly and regularly drawn on such resources), but rather through the hard and contingent work demanded in complex and contentious political settings. In this regard, the exigencies of the political field brought out, exploited, and in some measure created latent instabilities in the technical field. But once opened, such controversies were not easily or quickly resolved, precisely because of their tendency to spill across the conjoined worlds of technical and political action.

Second, despite the arguably distinctive intensity of the California water case, the general presence of simulation tools at the center of contested domains of public policy seems unlikely to decrease in future (though given the challenges and instabilities noted above, this is not a foregone conclusion). The rapid and comparatively recent development of simulation techniques and their generally speedy diffusion through the policy field to date would seem to suggest that 'model knowledges' – conclusions, predictions, and other assertions of fact drawn partly or primarily on the basis of computer simulations – are likely to figure as more rather than less significant policy technologies in future. Under such circumstances, the sorts of model literacy advocated above may become an increasingly important attribute and skill-set, both within digital government scholarship and democratic polities more generally.

Third, as the case study suggests, the work required to build and sustain models as meaningful objects in the world comes in many forms: conceptual, mathematical, and computer-based; but also organizational, political, and broadly sociological. The latter is sometimes treated as an add-on to the real work of modeling, perhaps necessary but fundamentally distinct from the technical work of building and running the 'models themselves' (a form of what we might call 'code realism'). In the world of California water modeling, however, such distinctions are hard and arguably becoming harder to maintain. As the field study traced above will begin to suggest, it turns out to be extraordinarily difficult to assign where, precisely, the work of modeling begins and ends. Can a model be reduced to code? To data? To the immediate network of designers and decision-makers that build and use it? To the wider networks of trust and credibility that sink or sustain its claims? The diversity of ecologies within which models build and hold meaning give them, like other many other complex artifacts, a considerable degree of 'ontological sprawl' that is neither easily nor obviously reduced. Effective modeling, like other instances of technical work in the policy arena, must attend to this diversity by meeting head-on the full range of technical, institutional, and broadly sociological conditions that enable and constrain its work.

6. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 0131923, and by NSF Dissertation Improvement Grant No. 0425261. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. The author wishes to thank Geoffrey Bowker, Robert Horwitz, Jane Fountain, David Lazer, and Sheila Jasanoff for guidance and support during the research and writing of this project.

7. REFERENCES

- [1] Beven, K. Changing Ideas in Hydrology: The Case of Physically-Based Models. *Journal of Hydrology* 105 (1989), 157-172.
- [2] Beven, K. Prophecy, reality and uncertainty in distributed hydrological modeling. *Advances in Water Resources* 16 (1993), 41-51.
- [3] California Department of Water Resources. *The State Water Project Delivery Reliability Report*. Sacramento, Aug 2002.
- [4] California Department of Water Resources. *CalSim II Simulation of Historical SWP-CVP Operations: Technical Memorandum Report*. Sacramento, Aug 2003.
- [5] California Water Plan Advisory Committee Modeling Work Group, Update 2003 Modeling Proposal Addendum (email to Advisory Committee). Accessed Aug 9 2005 at: <http://www.waterplan.water.ca.gov>
- [6] Canham, C., Cole, J., and Lauenroth, W. Models in Ecosystem Science. In *Models in Ecosystem Science*, eds. Canham, C., Cole, J., and Lauenroth, W. Princeton University Press, Princeton, 2003, 1-12.

- [7] Carpenter, S. The Need for Fast-and-Frugal Models. In *Models in Ecosystem Science*, eds. Canham, C., Cole, J., and Lauenroth, W. Princeton University Press, Princeton, 2003, 455-460.
- [8] Close, A., Haneman, W., Labadie, J., Loucks, D., Lund, J., McKinney, D., Stedinger, J. A Strategic Review of CALSIM II and its Use for Water Planning, Management, and Operations in Central California. Panel report submitted to the California Bay Delta Authority Science Program Dec 4, 2003.
- [9] Collins, H. *Changing Order: Replication and Induction in Scientific Practice*. University of Chicago Press, Chicago, 1985.
- [10] Cottingham, K. et. al. Increasing Modeling Savvy: Strategies to Advance Quantitative Modeling Skills for Professionals within Ecology. In *Models in Ecosystem Science*, eds. Canham, C., Cole, J., and Lauenroth, W. Princeton University Press, Princeton, 2003, 428-436.
- [11] Draper, A., Munevar, A., Arora, S., Reyes, E., Parker, N., Chung, F., Peterson, L. CalSim: A Generalized Model for Reservoir System Analysis. Unpublished paper available at: <http://science.calwater.ca.gov/index.html>.
- [12] Edwards, P. Global Climate Science, Uncertainty, and Politics: Data-Laden Models, Model-Filtered Data. *Science as Culture* 8:4 (1999), 437-472.
- [13] Ezrahi, Y. *The Descent of Icarus: Science and the Transformation of Contemporary Democracy*. Harvard University Press, Cambridge, MA, 1990.
- [14] Ferreira, I., Tanaka, S., Hollinshead, S., and Lund, J. *CALSIM II in California's Water Community: Musing on a Model*. Report prepared for the CALFED Science Program, Jan 20, 2004.
- [15] Fox-Keller, E. Models of and models for: theory and practice in contemporary biology. *Philosophy of Science* 67 (2002), S72-82.
- [16] Galison, P. *Image and Logic: A Material Culture of Microphysics*. University of Chicago Press, Chicago, 1997.
- [17] Gusterson, H. *People of the Bomb: Portraits of America's Nuclear Complex*. University of Minnesota Press, Minneapolis, 2004.
- [18] Jackson, S. *Building the Virtual River: Numbers, Models, and the Politics of Water*. Unpublished Ph.D. Dissertation, University of California, San Diego, 2005.
- [19] Jasanoff, S. Technologies of Humility: Citizen Participation in Governing Science. *Minerva* 41 (2003), 223-244.
- [20] Jasanoff, S. *Designs on Nature: Science and Democracy in Europe and the United States*. Princeton University Press, Princeton, 2005.
- [21] Kiparsky, M., and Gleick, P. Climate Change and California Water Resources: A Survey and Summary of the Literature. Report by Pacific Institute for Studies in Development, Environment, and Security, July 2003.
- [22] Kraemer, K., Dickhoven, S., Tierney, S. and King J. *Datawars: The Politics of Modeling in Federal Policymaking*. Columbia University Press, New York, 1987.
- [23] Kwa, C. Modeling Technologies of Control. *Science as Culture* 4:20 (1994), 363-391.
- [24] Lauenroth, W., Burke, I., and Berry, J. The Status of Dynamic Quantitative Modeling in Ecology. In *Models in Ecosystem Science*, eds. Canham, C., Cole, J., and Lauenroth, W. Princeton University Press, Princeton, 2003, 32-48.
- [25] Lund, J. In the dark, all cats are gray (letter to the Department of Water Resources, Division of Planning and Local Assistance, Sep. 6, 2003).
- [26] O'Connor, D. Comments on the Department of Water Resources' Draft State Water Project Supply Reliability Report, Nov 1, 2002.
- [27] Oreskes, N., Shrader-Frechette, K., and Belitz, K. Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science* 263:5147 (Feb. 4, 1994): 641-46.
- [28] Oreskes, N., and Belitz, K. Philosophical Issues in Model Assessment. In *Model Validation: Perspectives in Hydrological Science*, eds. Anderson, M. and Bates, P. John Wiley and Sons, New York, 2001, 23-42.
- [29] Porter, T. *Trust In Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton University Press, Princeton, 1995.
- [30] Shackley, S. and Wynne, B. Global Climate Change: The Mutual Construction of an Emergent Science-Policy Domain. *Science and Public Policy* 22:4 (1995), 218-230.
- [31] Shackley, S. and Wynne, B. Representing Uncertainty in Global Climate Change Science and Policy: Boundary-Ordering Devices and Authority. *Science, Technology, and Human Values* 21:3 (1996), 275-302.
- [32] Sismondo, S. Models, Simulations, and Their Objects. *Science in Context* 12 (1999), 247-260.
- [33] Star, S.L., Grisemer, J. Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science* 19 (1989), 387-420.
- [34] Thevenot, L., Boltanski, L. *On Justification: Economies of Worth* (tr. Catherine Porter). Princeton University Press, Princeton, 2006.
- [35] Wilkinson, R. Comments on DWR's Draft Report 'The State Water Project Delivery Reliability Report', Oct 31, 2002.
- [36] Wolff, G., and Gleick, P. The Soft Path for Water. *The World's Water 2002-2003: The Biennial Report on Freshwater Resources*. Island Press, Washington, 2002, 1-32.